

WebMedline: Transforming Medline into a Hypertext Environment with Links to Full-Text Documents

William M. Detmer and Edward H. Shortliffe

Section on Medical Informatics, Stanford University School of Medicine

<http://smi.stanford.edu/people/bdetmer>

Many studies have shown that the daily information needs of clinicians are seldom met. A major barrier to meeting these needs is timely access to current, relevant knowledge. Progress could be made if resources such as Medline, evidence-based-medicine reviews, and full-text articles could be seamlessly integrated and made available to clinicians across computer networks.

METHODS

To explore this vision, we have developed WebMedline (1), a World Wide Web Common Gateway Interface (CGI) application that transforms the Medline database into a hypertext environment. WebMedline receives user requests in the form of Hypertext Markup Language (HTML) forms, queries a remote Medline database, finds full-text-document links relevant to retrieved Medline citations, and then displays the citations and links in hypertext.

WebMedline uses the Melvyl Medline database maintained by the University of California. The standard method of accessing Melvyl Medline is to initiate a terminal-based Telnet session, to logon to the Melvyl host, to navigate the opening prompts, to issue queries in the Melvyl Medline query language, and to display results using a specialized language.

In contrast, WebMedline users enter text into predefined fields, such as Author, Title, Journal, and Keyword. Using pop-up menus, they can also choose the database years, the between-field Boolean operator, the display type, and the number of citations to retrieve. Finally, they can choose to constrain the search by standard limiters such as "English," "Human subjects," and "Publication Type" (e.g., randomized controlled trial).

When the user submits a search, the contents of the HTML form are first submitted to a Hypertext Transfer Protocol (HTTP) server and then passed to the WebMedline CGI. The CGI first formulates a legal Melvyl Medline query by removing stop words, stemming words to their roots, and appending field descriptors. The CGI then initiates a Telnet session with the remote Melvyl Medline host, logs on using a password supplied by the user, navigates the opening prompts, issues the fully specified Melvyl query, and retrieves the specified number of citations in the desired format. The CGI then iterates through the citations: first identifying the NLM unique identifier associated with each citation, then querying the WebMedline link database for full-text documents

associated with the unique identifier, and finally marking up the citation with hypertext links to the full Medline record and any full-text documents.

The WebMedline link database consists of a relational table with entries for each full-text document related to a Medline citation. Fields in this database include: NLM unique identifier, Uniform Resource Locator (URL) of document, document type (e.g., full-text, evidence-based medicine review), document format (e.g., HTML), and access restriction (domain name of restricted access area, if any). Records in the database are generated by automated scripts that parse document collections for their source information and query Medline for corresponding unique identifiers.

RESULTS

WebMedline has been used at Stanford and UC-San Francisco since February 1995, and at UC-Davis since April, 1996. Currently, WebMedline links Medline citations to *ACP Journal Club* reviews, NIH Consensus Development Conferences, and full-text articles from the *Journal of Biological Chemistry*. Usage logs show that WebMedline has been used over 200,000 times and that usage is doubling every 4-6 months. As a percentage of all searching done on the Stanford campus, WebMedline searches continues to grow from 5% in July 1995 to nearly 30% in June 1996.

CONCLUSIONS

WebMedline demonstrates how "middleware" applications on the World Wide Web can integrate and display data from disparate information sources. Future work will focus on methods for robust client-server interaction between Web browsers and information sources and evaluation of the impact of such systems on clinical information needs.

Acknowledgments

This work was conducted with the support of the National Library of Medicine under grant LM-07033. Computing facilities were provided by the CAMIS Resource, LM-05305. The *ACP Journal Club* reviews were generously provided by the American College of Physicians.

Reference

1. Detmer WM, Shortliffe EH. A model of clinical query management that supports integration of biomedical information over the World Wide Web. *Proc Annu Symp Comput Appl Med Care*. 1995:898.